A Web app for document analysis, research and note-taking

STUDY

INTRODUCTION

- My name is Saikat Mukherjee. I am a software developer currently working in Amazon.

- I Completed my Bachelor's from RCC Institute of Information Technology in India in 2016 and worked for 4 years in Tata Consultancy Services. I came to US in 2021 to Pursue Masters degree

- I completed my Masters in Robotics from Arizona State University in 2022 and joined Amazon in Bellevue, Washington.

- I am a self-taught Machine Learning Engineer who doesn't have traditional work experience in this field.

- I built ML/DL models using TesnsorFlow to detect object in live video stream, developed models to classify Images and Texts

- Currently I am building apps using Generative Machine Learning APIs from Open AI and more.

PROBLEM STATEMENT

- How about having a web app that can answer questions from any PDF document using natural language processing not only in text but using voice.

- But what if the user wants to know more about a specific topic from that pdf and search for current information or information that is not the pdf , from internet.

- The icing of the cake would be if the user can take notes during this information retrieval process and download that file for future reference after he/she is done.

- Moreover, if the app has all these along with a separate page where the user can ask any questions to ChatGPT, wouldn't that be great?

SOLUTION

- In this app the user will be able to upload any pdf and will able to retrieve information by asking questions in chat format.

- The answers will be available in audio format as well which will increase the accessibility of the application.

- In case the user wants to search more about a particular topic from pdf, beyond the scope of the literature present in PDF, he/she will be able to do that as well in a separate page called **Research.** Where answers will be retrieved from internet through Google Search and Wikipedia.

- There is a separate page to takes notes during these learning process where the text **persists** even if the user goes to a different page and at the end user will be able to download it.

- Last but not least, There is a separate page where User will be able to ask questions to chatGPT.

TECHNICAL DETAILS

- Used Streamlit to create the web app.

- Used ChatGPT GPT3.5 API for llm in LangChain

- Used PyPDFLoader, ChromaDB, OpenAIEmbeddings to retrieve information from PDF for Q & A

- Used Text-To-Speech API from Google Cloud to get answers from PDF in Audio along with text.

- Used LangChai Tools for i.e. Wikipedia, SerpAPI for research page.

Future Work

- The App can be further enhanced by making it fully driven by voice command.

- Other than PDF, different types of files can be supported for information retrieval

- Making the information retrieval in multiple languages would decrease the language barrier

- Transcribing any Audio or video file from URL would make the app more robust.